

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/940,188	08/27/2001	Dongming Jiang	25961-706	3518
21971	7590	09/03/2004	EXAMINER	
WILSON SONSINI GOODRICH & ROSATI 650 PAGE MILL ROAD PALO ALTO, CA 943041050			CHEN, TE Y	
			ART UNIT	PAPER NUMBER
			2171	

DATE MAILED: 09/03/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	09/940,188	JIANG ET AL.
	Examiner	Art Unit
	Susan Y Chen	2171

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 04/4/2004 & 06/2/2004.
- 2a) This action is **FINAL**. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-23 and 25 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 1-23 &25 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All
 - b) Some *
 - c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.



Attachment(s)

- 1) Notice of References Cited (PTO-892)
- 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____.
- 4) Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
- 5) Notice of Informal Patent Application (PTO-152)
- 6) Other: _____.

Response to Amendment

This office action is in response to amendments filed on April 6, 2004 and June 2, 2004. Claims 1, 7 and 12 have been amended and claim 24 has been canceled at April 6, 2004. Claim 25 has been newly added on June 2, 2004.

Specification

The disclosure is objected to because of the following informalities:

In previous office action (filed on 10/06/2003) the examiner specifically request that applicant to check the extent of the disclosure for determining the presence of all possible minor errors, however, applicant fails to do so, for example, the cross-reference of the instant application still contains unfilled serial number.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before

the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims 1-23 and 25 are rejected under 35 U.S.C. 102(e) as being anticipated by Brady et al. (U.S. Patent No. 6,463,430).

As to claims 1, 3, 5, 10 and 25, Brady et al. (hereinafter referred as Brady) discloses the means and methods of focused crawling system, comprising:

a) accessing a query input [e.g. the Front-End component (108), Fig. 1; col. 15, lines 14-19; Fig. 8];

b) crawling a plurality of documents via the guidance of a crawl metric [e.g., the Retrieval Priority List, col. 4, line 66 – col. 5, line 1; col. 5, lines 9-11], wherein the crawl metric is partially determined by a first mechanism [e.g. the ranking mechanism applied to the content of Retrieved Priority List; col. 5, lines 3-5] via the steps of: 1) evaluating the relevance of documents using logical expression of keywords and phrases [e.g., the subject taxonomy processing applied to the example documents stored in the retrieval priority list, col. 5, lines 41-47; Fig. 11]; 2) evaluating relevance of documents using a template [e.g. a Web page template, col. 4, lines 1-11] which contains at least one template portion and the portion contains at least one level [e.g. the fields in the document, col. 4, lines 30-32]; 3) evaluating relevance of documents using a link structure of the crawled documents [e.g. the Retrieval Priority List is a linked structure, col. 4, line 66 – col. 5, line 1; col. 5, lines 9-118]; 4) evaluating

relevance based on freshness of documents [e.g., the Retrieval Priority List is repeatedly updated with fresh documents (i.e. the links) being added and the content of the Retrieval Priority List are repeatedly ranked according to their relevance to a subject taxonomy. Col. 5, lines 3-11; col. 8, lines 22-46; Fig. 5].

c) returning target documents based on a search metric which is determined by a second mechanism which repeats the steps of 1-4 of the first mechanism [e.g., see Fig. 11, the outer-loop processing mechanism (1106-1120) of unit 1112]. Wherein, the procedure, of the first mechanism, evaluating relevance of documents using a link structure of the crawled documents [e.g., the Retrieval Priority List (1114) is used by the inner-loop mechanism (1108-1116), Fig. 11]. In addition, the crawl metric comprises a metric that quantifies priority for crawling links emanating from a certain document within context of the focused crawling [e.g., the extracted resource links are place in the processing queue of the Data store ranked by the score, 806, Fig.8; col. 19, lines 7-13]. And the search metric comprises a metric that quantifies relevance of a document to the query input [e.g., the score page processing, 804-814, Fig. 8];

d) accessing a first plurality of documents from a database of documents , wherein the documents including crawled documents which are ranked [e.g., col. 9, lines 10-42];

e) generating a graph of the first plurality of documents [e.g. the Taxonomy section (1301, Fig. 13) can be created via the creating a new directory or creating a new resource functions, col. 24, lines 5-19; col. 23, lines 32-40; col. 24, line 55 – col. 25, line 25];

- f) assigning weights to one or more nodes of the graph [e.g., the Score Page and Score links unit (806, Fig. 8)];
- g) finding an assignment of weights to one or more modes of the graph, by propagating weights through the graph, the assignment of weight to a node partially based on calculating a weighted sum of weights propagated from neighboring nodes [e.g., the steps 804-812, Fig. 8];
- h) generating a ranked list which is partly generated from the graph [e.g., the steps 1106-1118, Fig. 11];

As to claims 2, 4, 6 and 11, the claimed feature – relevance includes importance is defaulted by the page ranking technique.

As to claims 7 and 12, Brady further discloses that the first mechanism including: a) associates a weight to each of the evaluated relevance of the procedures [e.g. col. 22, lines 1-14]; b) combining the evaluated relevance and the weights of the evaluated relevance [col. 21, lines 56-65].

As to claims 8 and 13, Brady further discloses that the using of template for evaluating relevance includes at least one heading levels and content levels evaluating processing [e.g., col. 22, lines 44 - 46].

As to claims 9 and 14, Brady further discloses that the evaluation relevance includes evaluating relevance of a first document [e.g. a page] and the

link which being used to refer the second plurality sets of documents [col. 21, line 37 – col. 22, line 29].

As to claims 15-20, Brady further discloses that the evaluation relevance processing can be done by expanding or shrinking the graph with a second set of one or more documents from the database via links of the document, in a forward and backward direction to produce the third set of document as claimed [col. 18, lines 32-48; the steps of 1301-1303, Fig. 13].

As to claim 21, Brady further discloses that the evaluation combining process is based on common characteristics of the nodes or relationships between the nodes [e.g. col. 21, lines 61-65].

As to claim 22, Brady further discloses that the weight propagating processing is done by a limited node distance [e.g. the limited threshold value, col. 22, lines 25-27].

As to claim 23, Brady further discloses that the weight assigned to a document includes at least one of the relevance of the document to the query input and importance of the document independent of the query input [e.g., col. 21, line 59 – col. 22, line 6].

Response to Arguments

Applicant's arguments filed on April 6, 2004 have been fully considered but they are not persuasive.

The examiner disagrees with applicant's argument that "Brady fails to teach crawling at least partially guide by a crawl metric, and documents being returned at least partially based on a search metric".

In reply to this argument, the examiner points out that Brady specifically discloses the claimed features at col. 5, lines 3-11, wherein, Brady clearly teaches using the retrieval priority list to direct (or guide) the Spider (or crawler) to retrieve the most relevant document identified by the process at any given moment. Furthermore, Brady specifically teaches that the documents being returned at least based on a search metric [e.g., the categorized scored pages, the steps 804-814, Fig. 8].

The examiner further disagrees with applicant's argument that "the query input in claim 1 potentially includes input for both crawl metric and search metric, which are separate as discussed above. ...The Office Action has not shown any teaching in Brady of using such parts of a query for a separate crawl metric and search metric".

In reply to this argument, the examiner points out that applicant defines none of the structures for the claimed first and second query parts. In addition, the applicant also fails to specify the links between the claimed "a mechanism",

“the first query part”, “the target documents” and “the second query part”, therefore, the claims are open for any reasonable interpretation. As discussed in the above USC 102(e) rejection, the examiner believes that the logic flows and text depictions of Fig(s) 8 and 11 read the claimed features.

In addition, the examiner disagrees with applicant’s argument that “Brady does not teach generating a graph of the documents”.

In reply to this argument, the examiner first point out that applicant fails to defined the metes and bounds of claimed documents, thus, it is open for interpretation. Furthermore, the examiner directs applicant’s attention to the description of the taxonomy section [1301, Fig 13] at col. 24, line 55 – col. 25 line 25, wherein, Brady clearly teaches that the taxonomy can provides a graphical and textual representation of the taxonomy of the information contained in a resource database.

Base on the above discussions, the examiner maintains the same rejection.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is

filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Susan Y Chen whose telephone number is (703) 308-1155. The examiner can normally be reached on Monday - Friday from 7:00-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Safet Metjahić can be reached on (703) 308-1436. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Susan Y Chen
Examiner
Art Unit 2171

August 31, 2004



UYEN LE
PRIMARY EXAMINER